Aug /807
Paul Seg.
Adams
5/1/98

CRF Diskette Problem Report

The Scientific and Technical Information Center (STIC) experienced a problem when processing the following CRF diskette:

Application Serial Number: 08/530,112

Filing Date: 7/10/95

Classification: 536

Date Reviewed by STIC: 10/23/96

Point-of-Contact / Telephone No: __Anne-Marie Corrigan__
__703-308-4222__

Nature of Problem:

☐ The CRF diskette was:
 ☐ Damaged
 ☐ Unreadable
 ☐ Blank (no files present on the floppy disk)

☐ A computer virus was detected on the diskette. The STIC will not process the diskette through the Data Capture System.

 Name of the virus: _____

☑ The CRF diskette contains an error that disrupts normal processing, as explained below:

 ☐ The Sequence Listing was not converted into ASCII (DOS) text

 ☐ See attached pages for clarification --->
 ☑ Other: _see attached_

9/7/95

(2) INFORMATION FOR SEQ ID █████ 1: *Sequences ████ed*
 (i) SEQUENCE CHARACTERISTI███
~~SEQ ID NO:1~~ *(delete)*
(A) LENGTH:704
(B) TYPE:nucleic acid → (C) STRANDEDNESS: *mandatory response for nucleic acids*
(D) TOPOLOGY:linear → (vii) IMMEDIATE SOURCE:
(E) CLONE:HUMGS00001
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1

```
GATCTTCAAA CAAGCATCAG CGTTTTCCAG GGCTTCCCAG AGGTCTGTGC GACTAGCCCG  60
TGTCTATCAA AAGTTATTAG AGAGGATGAA GCATTAGCTT GAAGCACTAC AGGAGGAATG 120
CACCACGGCA GCTCTCCGCC AATTTCTCTC AGATTTCCAC AGAGACTGTT TGAATGTTTT 180
CAAAACCAAG TATCACACTT TAATGTACAT GGGCCGCACC ATAATGAGAT GTGAGCCTTG 240
TGCATGTGGG GGAGGAGGGA GAGAGATGTA CTTTTTAAAT CATGTTCCCC CTAAACATGG 300
CTGTTAACCC ACTTGCATGC AGAAACTTGG GATGTCACTT GCCTGACATT CACTTTCCAG 360
GAGAGGACCC TATCCCCAAA TGTGGAATTG ACTTGCCTAT GGCCAAGGTC CCTTGGNAAA 420
GGGAGCTTCA GTATTTGTGG GGGCNTCATA AAACCATGGN TTCAAGNCAA TCCAGCCTCA 480
TNGGGNNGGT CCTGGGNACA GTTTTTTGGT AAAGGCCCTT GGCCCAGNTG GGGGGAATGG 540
GCCTCCTTTT TAAGNTTTGG GNTGGAATNG TCTNGCAAAT TGGGGCTCCC ATTTCNCGGG 600
GGTTTGGGGG TTTTTTNGGG CCTTNCCNGG NNGGAAGGGN TGGGTTTGGG GGNTNGGTTN 660
CCNTTGGGNG GGCCTGGGGN TTTGATTTNA CCCGGGNCTT NGGN                  704
```

SEQ ID NO:2
LENGTH:659
TYPE:nucleic acid
TOPOLOGY:linear
CLONE:HUMGS00002
SEQUENCE DESCRIPTION:

use above headings

```
GATCTTTAAA ATACACACTC AAATCAAGAA ACTTAAGGTT ACCTTTNTTC CCAAATTTCA  60
TACCTATCAT CTTAAGTAGG GACTTCTGTC TTCACAACAN ATTATNACCT TACAGAAGTT 120
TGAATTATCC GGTCGGGTTT TATTGTTTAA AATCATTTCT GCATCAGCTG CTGAAACAAC 180
AAATAGGAAT TGTTTTTATG GAGGCTTTGC ATAGATTCCC TGAGCAGGAT TTTAATCTTT 240
TNCTAACTGG ACTGGTTCAA ATGTTGTNCT CTTCTTTAAA GGGATGGCAA GATGTGGGCA 300
GTGATGTCAC TTAGGGCAGG GACAGGATAA GAGGGNTTAG GGAGAGAAGA TAGCAGGGCA 360
TGGCTGGGAA CCCAAGTCCA AGCATACCAA CACGGAGCAG GCTACTGTCA AGCTCCCCTC 420
GGAGGCGGNG CTGGTTCACA GCCAGCTGGC ACCAGNTTTT NTNGNGGAAG NCTTTTTCAA 480
ACAGTCTCAG GNAATCCAAT NTGCAAAGAC TTGCTTTNAG NAAAACCCAG NAGTTGAAAG 540
GCTCCCAAGN ATTTTAAGGG NACTTNCCAA AACGGGGCCC CNGGNNCCTT TTGGGTTTNG 600
GGGNTCAAAA CCCCGGAGGG GTTTGGGAAG NTTTTAATTG GNTTTAAAAN ATNNNTNTN  659
```

SEQ ID NO:3
LENGTH:625
TYPE:nucleic acid
TOPOLOGY:linear
CLONE:HUMGS00003
SEQUENCE DESCRIPTION:

```
GATCTAACTG GGTACCTGAG ATATTTNACA GCTGGACCTA GTTTCACAAT CTGTTGTCTC  60
CAGCTCTGCA TATGTCTGGC CAGGGGGCTT CTAGGAAGTA GGTTTCATCT ATCAAATGTC 120
TCCTCTGACT TCCTTTTGAA ACTTACTGCT CTTCTGTTTT ATTTTGTTTT GTTTGAAGCT 180
CAGAGGGAGA TGGGCAATTG ACAGGGATGC AATCCAGGGT GGGATTTCTT GAGGAAGTTA 240
CAAATAAGCT TGTTACAACA TCAAGATAGA TGGAATTGGA AGGATGCTAC CAGGAGAGTA 300
CTTACATAGT GCTCAGGAGT TTCTCTTCTT AAAATGTTTA CTGCTGAAAG ATGAGCAGGA 360
CCAGGGCGTT ATAGGCAGAG CCCTAGCCGA GAAACCTGCT GGCCTCTGCC TGTTTTCATT 420
TCCCACTTTT GGTTGTTGTG GCATTACTTT CAGAATTTGC ACTTTCCTGC TTGTCATGAC 480
TTTTTTGGCA CACTTGCCAT GACGGGTGTT TCTGNGAACC ATGGAAGTTT TGCGGTAGTG 540
CCTCCAGGGG CAGGGGGNAA GGAGGNGGTG TANCTGCATT TNGTNCAAAT AAATCCNGCC 600
TATTGTTAAT NAACCAGTCT TTTGN                                       625
```

SEQ ID NO:4

(3) Computer: Apple Macintosh;

(i) Operating System: Macintosh;

(ii) Macintosh File Type: t xt with line termination

(iii) Line Terminator: Pre-defined by text type file;

(iv) Pagination: Pre-defined by text type file;

(v) End-of-file: Pre-defined by text type file;

(vi) Media: (A) Diskett—3.50 inch, 400 Kb storage;

(B) Diskette—3.50 inch, 800 Kb st rage;

(C) Diskette—3.50 inch, 1.4 Mb st rage;

(vii) Print Command: Use PRINT command from any Macintosh Application that processes text files, such as MacWrite or Teach Text;

(4) Magnetic tape: 0.5 inch, up to 2400 feet;

(i) Density: 1600 or 6250 bits per inch, 9 track;

(ii) Format: raw, unblocked;

(iii) Line Terminator: ASCII Carriage Return plus optional ASCII Line Feed;

(iv) Pagination: ASCII Form Feed or Series of Line Terminators;

(v) Print Command (Unix shell version given here as sample response—mt/dev/rmt0; lpr/dev/rmt0):

(g) Computer readable forms that are submitted to the Office will not be returned to the applicant.

(h) All computer readable forms shall have a label permanently affixed thereto on which has been hand printed or typed, a description of the format of the computer readable form as well as the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form and the name and type of computer and operating system which generated the files on the computer readable form. If all of this information cannot be printed on a label affixed to the computer readable form, by reason of size or otherwise, the label shall include the name of the applicant and the title of the invention and a reference number, and the additional information may be provided on a container for the computer readable form with the name of the applicant, the title of the invention, the reference number and the additional information affixed to the container. If the computer readable form is submitted after the date of filing

und r 35 U.S.C. 111, after the date of ntry in the national stage under 35 U.S.C. 371 or after the time of filing, in the United States Receiving Office, an international application under the PCT, the labels mentioned herein must also include the date of the application and the application number, including series code and serial number.

§ 1.825 Amendments to or replacement of sequence listing and computer readable copy thereof.

(a) Any amendment to the paper copy of the "Sequence Listing" (§ 1.821(c)) must be made by the submission of substitute sheets. Amendments must be accompanied by a statement that indicates support for the amendment in the application, as filed, and a statement that the substitute sheets include no new matter. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(b) Any amendment to the paper copy of the "Sequence Listing," in accordance with paragraph (a) of this section, must be accompanied by a substitute copy of the computer readable form (§ 1.821(e)) including all previously submitted data with the amendment incorporated therein, accompanied by a statement that the copy in computer readable form is the same as the substitute copy of the "Sequence Listing." Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(c) Any appropriate amendments to the "Sequence Listing" in a patent, e.g., by reason of reissue or certificate of correction, must comply with the requirements of paragraphs (a) and (b) of this section.

(d) If, upon receipt, the computer readable form is found to be damaged or unreadable, applicant must provide, within such time as set by the Commissioner, a substitute copy of the data in computer readable form accompanied by a statement that the substitute data is identical to that originally filed. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

Appendix A—Sample Sequence Listing

(1) GENERAL INFORMATION:

(i) APPLICANT: Doe, Joan X, Doe, John Q

(ii) TITLE OF INVENTION: Isolation and Characterization of a Gene Encoding a Protease from *Paramecium* sp.

(iii) NUMBER OF SEQUENCES: 2

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Smith and Jones

(B) STREET: 123 Main Street

(C) CITY: Smalltown

(D) STATE: Anystate

(E) COUNTRY: USA

(F) ZIP: 12345

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage

(B) COMPUTER: Apple Macintosh

(C) OPERATING SYSTEM: Mcintosh 5.0

(D) SOFTWARE: MacWrite

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/999,999

(B) FILING DATE: 28-FEB-1989

(C) CLASSIFICATION: 999/99

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/US88/99999

(B) FILING DATE: 01-MAR-1988

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Smith, John A.

(B) REGISTRATION NUMBER: 00001

(C) REFERENCE/DOCKET NUMBER: 01-0001

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (909) 999-0001

(B) TELEFAX: (909) 999-0002

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 954 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(iii) HYPOTHETICAL: yes

(iv) ANTI-SENSE: no

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Paramecium sp

(C) INDIVIDUAL/ISOLATE: XYZ2

(G) CELL TYPE: unicellular organism

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: genomic

(B) CLONE: Para-XYZ2/36

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Doe, Joan X, Doe, John Q

(B) TITLE: Isolation and Characterization of a Gene Encoding a Protease from *Paramecium* sp.

(C) JOURNAL: Fictional Genes

(D) VOLUME: I

(E) ISSUE: 1

(F) PAGES: 1-20

(G) DATE: 02-MAR-1988

(K) RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 954

BILLING CODE 3510-16-M

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATCGGGATAG TACTGGTCAA GACCGGTGGA CACCGGTTAA CCCCGGTTAA GTACCGGTTA   60

TAGGCCATTT CAGGCCAAAT GTGCCCAACT ACGCCAATTG TTTTGCCAAC GGCCAACGTT  120

ACGTTCGTAC GCACGTATGT ACCTAGGTAC TTACGGACGT GACTACGGAC ACTTCCGTAC  180

GTACGTACGT TTACGTACCC ATCCCAACGT AACCACAGTG TGGTCGCAGT GTCCCAGTGT  240

ACACAGACTG CCAGACATTC TTCACAGACA CCCC ATG ACA CCA CCT GAA CGT CTC  295
                                    Met Thr Pro Pro Glu Arg Leu
                                                          -30

TTC CTC CCA AGG GTG TGT GGC ACC ACC CTA CAC CTC CTC CTT CTG GGG   343
Phe Leu Pro Arg Val Cys Gly Thr Thr Leu His Leu Leu Leu Leu Gly
        -25                     -20                 -15

CTG CTG CTG GTT CTG CTG CCT GGG GCC CAT GTGAGGCAGC AGGAGAATGG      393
Leu Leu Leu Val Leu Leu Pro Gly Ala His
      -10                     -5

GGTGGCTCAG CCAAACCTTG AGCCCTAGAG CCCCCCTCAA CTCTGTTCTC CTAG GGG    450
                                                            Gly

CTC ATG CAT CTT GCC CAC AGC AAC CTC AAA CCT GCT GCT CAC CTC ATT   498
Leu Met His Leu Ala His Ser Asn Leu Lys Pro Ala Ala His Leu Ile
  1           5               10                  15

GTAAACATCC ACCTGACCTC CCAGACATGT CCCCACCAGC TCTCCTCCTA CCCCTGCCTC  558

AGGAACCCAA GCATCCACCC CTCTCCCCCA ACTTCCCCCA CGCTAAAAAA AACAGAGGGA  618

GCCCACTCCT ATGCCTCCCC CTGCCATCCC CCAGGAACTC AGTTGTTCAG TGCCCACTTC  678

TAC CCC AGC AAG CAG AAC TCA CTG CTC TGG AGA GCA AAC ACG GAC CGT   726
Tyr Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg Ala Asn Thr Asp Arg
          20                  25                  30

GCC TTC CTC CAG GAT GGT TTC TCC TTG AGC AAC AAT TCT CTC CTG GTC   774
Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn Asn Ser Leu Leu Val
      35                  40                  45

TAGAAAAAAT AATTGATTTC AAGACCTTCT CCCCATTCTG CCTCCATTCT GACCATTTCA  834

GGGGTCGTCA CCACCTCTCC TTTGGCCATT CCAACAGCTC AAGTCTTCCC TGATCAAGTC  894

ACCGGAGCTT TCAAAGAAGG AATTCTAGGC ATCCCAGGGG ACCCACACCT CCCTGAACCA  954
```

(2) INFORMATION FOR SEQ ID NO: 2:
(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 82 amino acids
  (B) TYPE: amino acid
  (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(ix) FEATURE:
  (A) NAME/KEY: signal sequence
  (B) LOCATION: –34 to –1

(C) IDENTIFICATION METHOD: similarity
  to other signal sequences, hydrophobic
(D) OTHER INFORMATION: expresses
  protease
(x) PUBLICATION INFORMATION:
  (A) AUTHORS: Doe, Joan X, Doe, John Q
  (B) TITLE: Isolation and Characterization
    of a Gene Encoding a Protease from
    Paramecium sp.

(C) JOURNAL: Fictional Genes
(D) VOLUME: I
(E) ISSUE: 1
(F) PAGES: 1–20
(G) DATE: 02-MAR-1988
(K) RELEVANT RESIDUES IN SEQ ID NO:
  2: FROM –34 TO 48
BILLING CODE 3510-18-M

Here's where sequence 2 starts (after the sequence data of SEQ ID NO:1:)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Thr Pro Pro Glu Arg Leu Phe Leu Pro Arg Val Cys Gly Thr Thr
              -30                 -25                 -20

Leu His Leu Leu Leu Leu Gly Leu Leu Leu Val Leu Leu Pro Gly Ala
          -15                 -10                 -5

His Gly Leu Met His Leu Ala His Ser Asn Leu Lys Pro Ala Ala His
          1               5               10

Leu Ile Tyr Pro Ser Lys Gln Asn Ser Leu Leu Trp Arg Ala Asn Thr
15                  20                  25                  30

Asp Arg Ala Phe Leu Gln Asp Gly Phe Ser Leu Ser Asn Asn Ser Leu
              35                  40                  45

Leu Val
```

Notice of Availability

Applicant Aid for Biotechnology Computer Readable Form (CRF)
Sequence Listings Submissions

The Patent and Trademark Office (PTO) has developed a computer
program, called Checker, that will aid applicants in identifying
and correcting errors prior to making submissions for compliance
with the Requirements for Patent Applications Containing
Nucleotide Sequence and/or Amino Acid Sequence Disclosures
(sequence rules: 37 CFR 1.821 through 1.825). (Final rules were
published in the *Federal Register* (55 FR 18230) on May 1, 1990,
and in the PTO *Official Gazette* (1114 Off.Gaz.PatOffice 29) on
May 15, 1990.)

Checker is a DOS-based software program that is intended to
assist users in determining whether errors may be present in the
sequence listings, and is not intended to guarantee that the
submission is error-free.

The most current version of the software will be available via
computer downloading (details below). Copies on diskette are
also available. Updated software versions will not be
automatically mailed out; any updates will be announced in the
PTO *Official Gazette*.

The software can be accessed/requested in the following
locations:
1)   Dial-up access to the Patent and Trademark Office Bulletin
     Board System.
     Phone number: 703-305-8950
     Cost: Free-of-charge
2)   Dial-up access through the Internet. FTP site: ftp.uspto.gov
     Login as "anonymous". Software is in directory /pub/checker
     Cost: Free-of-charge
3)   For diskette copies, telephone requests to 703-306-2600.
     Cost: $25.00

For Further Information Contact: Meredith Beckhardt at 703-308-4212.